

Détection de regards de visite sur des images à haute résolution spatiale par une méthode d'apprentissage

Benjamin COMMANDRE^{1,2}, Driss EN-NEJJARY^{1,3}, Lionel PIBRE^{3,6}, Marc CHAUMONT⁴, Gérard SUBSOL⁵, Laurent DERUELLE⁶, Mustapha DERRAS⁶, Carole DELENNE^{1,2}, Nanée CHAHINIAN⁷

¹Univ. Montpellier, HydroSciences Montpellier, (UM / CNRS / IRD).

²INRIA Lemon, Montpellier.

³Univ. Montpellier, LIRMM, (UM / CNRS).

⁴Univ. Nîmes, LIRMM, (UM / CNRS).

⁵CNRS, LIRMM (UM / CNRS).

⁶BERGER-LEVRAULT, Recherche et Innovation, Toulouse-Labège.

⁷IRD, HydroSciences Montpellier, (UM / CNRS / IRD).

Mots clés : Réseaux de neurones; très haute résolution; cartographie; réseaux d'assainissement

L'utilisation de la télédétection en milieux urbains s'est focalisée dans un premier temps sur l'identification et l'estimation des surfaces urbaines (Weng, 2011), la mise en évidence de l'expansion urbaine (Xian et Crane, 2005), la quantification du microclimat (Arthur-Harthaft et al., 2001) ou de l'îlot de chaleur urbain (Dousset et Gourmelon, 2004; Gallo et al., 1994). En combinant les approches multi-spectrales et multi-temporelles, des suivis ont ainsi pu être faits sur l'évolution de ces phénomènes. A mesure que la résolution spatiale des images s'améliorait, l'identification d'objets précis du paysage urbain devenait possible (Bhaskaran et al., 2010). Ainsi, depuis quelques années, les images sont utilisées pour détecter les réseaux routiers (Péteri et Ranchin, 2006, Stoica et al. 2004), la végétation urbaine (Höfle, 2012), les caractéristiques du bâti (Hermosilia et al., 2011; Freire et al., 2010; Lhomme et al., 2004). En comparaison, la localisation d'objets de plus petite dimension (inférieure à un immeuble) reste relativement peu étudiée et pose le problème de l'identification d'objets dans un environnement variable, sur fond non-uniforme et souvent peu contrasté.

Dans le cadre du projet Cart'eaux*, nous proposons une méthode de détection de regards de visite (plaques d'égout) sur des images à très haute résolution spatiale. Cette méthode par apprentissage consiste à appliquer un réseau de neurones à convolution, AlexNet, sur des orthophotos à très haute résolution spatiale prises dans le visible.

Le réseau est entraîné sur une image à 5cm de résolution de la ville de Prades-Le-Lez dans l'Hérault. Afin de pallier le manque de données – seulement 605 plaques d'égout présentes sur l'image – une procédure d'augmentation de la base d'apprentissage est appliquée : rotation, miroir... conduisant à une base d'apprentissage à 2 classes constituée de 18405 « plaques d'égout » et 458915 « autres objets ». Une fois le réseau entraîné et « boosté », il est appliqué sur la ville de Gigean avec une précision de 72% et un rappel de 54%.

Plusieurs pistes de recherche sont actuellement explorées afin d'améliorer ces résultats déjà encourageants. D'une part, l'utilisation d'une information altimétrique (MNT, LIDAR) ou de données géographiques (bâti) permettrait de réduire le nombre de faux positifs (e.g. cheminées) ; d'autre part l'application d'un filtre géométrique en amont de la classification fournira une information supplémentaire au réseau de neurones qui devrait augmenter le nombre de vrais positifs.

* Projet financé par la région Occitanie et l'Europe (fond FEDER)